Molecular Biology III: Methods of Nucleic Acid Analysis

I. Learning Objectives.

A. For electrophoresis explain:

1. the physical basis.
2. the methods for detecting nucleic acids after electrophoresis.

B. Compare and contrast Southern, Northern, and Western blotting techniques.

C. Compare the different methods for identifying polymorphisms (RFLP, VNTR and SNP).

D. Describe the *polymerase chain reaction (PCR)* for the rapid amplification of specific DNA fragments. List three applications of PCR.

E. Explain the *Sanger dideoxy chain termination* method of DNA sequencing.

F. Textbook reading for reference and clarification. — Lehninger, Chapter 29; Gelehrter, Chapter 5.

II. Introduction.

The methods of nucleic acid analysis are used for distinguishing different DNA molecules on the basis of their different nucleotide sequences.

III. Polymerase Chain Reaction (PCR) – A single tube method for making numerous copies of DNA from a small amount of DNA.

A. DNA is copied (synthesized) using a heat-resistant DNA polymerase (obtained from bacteria that inhabit thermal vents).

B. Small amounts of DNA (from as little as a single cell) are amplified into a large number of copies (millions to billions).

C. Components of the PCR

1. *four nucleotide triphosphates* (dATP, dTTP, dCTP, dGTP)
2. *heat resistant DNA polymerase* – catalyzes the synthesis of new DNA.
3. *template DNA* – the DNA to be amplified.
4. *synthetic oligonucleotide "primers"* – Two primers are used, one complementary to each of the two strands of the *template* DNA. The polymerase catalyzes phosphodiester bond formation between the incoming nucleotide and the 3' hydroxyl group of the primer (nucleophilic attack by the 3'-hydroxyl oxygen of the 3' terminal nucleotide and the alpha phosphate of the incoming nucleotide triphosphate). The synthetic primers define the region that will be amplified. The choice of the "primers" determines the sequence that is amplified.

D. The reaction cycle (typically 20-35 cycles are used for an amplification)□. See figure on next page.

1. *Denaturation* of duplex "template" (target DNA) – achieved by brief heating to about 95˚C.

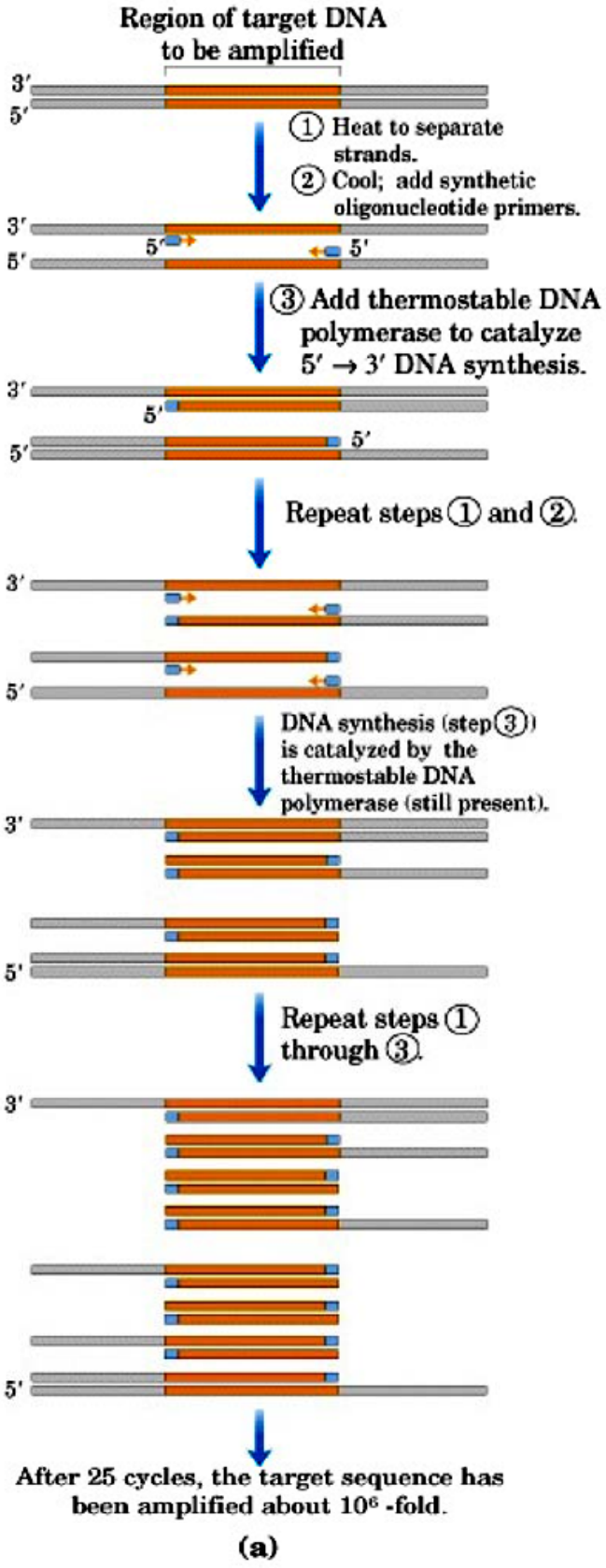
2. *Annealing* of primers to template – achieved by cooling to ~45-60˚C (about 5˚C below the T_m).
3. *Polymerization* of nucleotide triphosphates into the strand complementary to the template by the heat-resistant polymerase (~72˚C). This is also called the primer extension phase.
4. 25 to 30 cycles can take less than an hour. In each cycle, the DNA between the primers doubles and therefore, after 30 cycles, the DNA is amplified 2^{30} times. Thus millions to a billion or more copies can be obtained from only a few DNA strands.

E. PCR is extremely sensitive. It can detect a single gene in a single cell. Moreover, it is simple and rapid. PCR is now routinely used in clinical applications (i.e. detecting viral infections, prenatal and postnatal diagnosis of genetic diseases, clinical studies on the association of diseases with potential predisposing genes, forensics, biological weapons detection).

IV. Southern Blotting. A method for detecting specific restriction fragments of DNA present in a complex population of restriction fragments such as is produced when genomic DNA is cut with a restriction enzyme. When a complex genome such as the human genome is cut with a restriction enzyme, there are so many different sized fragments that an ethidium bromide-stained agarose gel containing these fragments looks like a continuous smear of DNA. It is therefore necessary to use a probe to "light up" (hybridize a labeled probe) the restriction fragment of interest in order to detect it. See figure on next page.

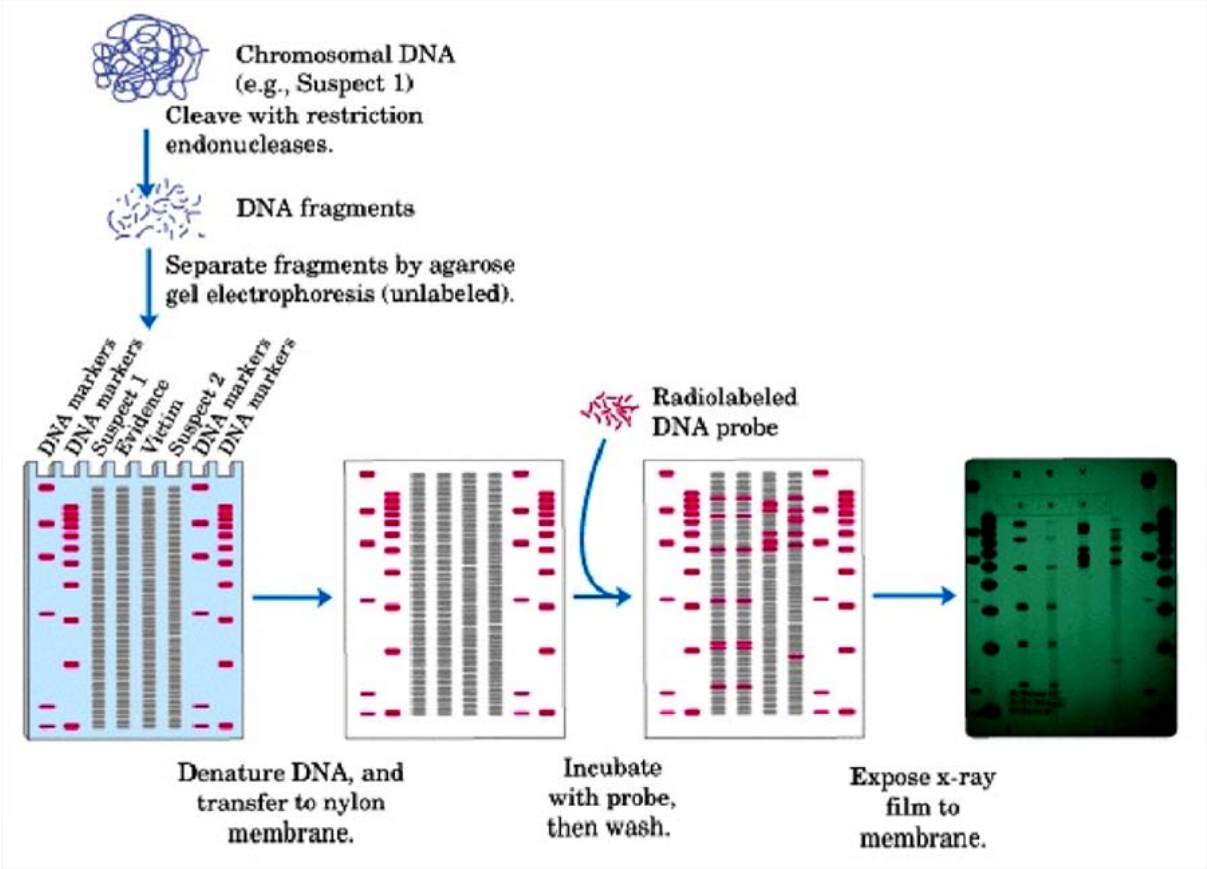
A. Procedure

1. Digest DNA with restriction



enzymes.

2. Electrophorese the resulting restriction fragments to separate them according to size.
3. Soak the gel in alkali to denature the duplex DNA into single strands (so that *probes* can hybridize in a later step).
4. Transfer the denatured DNA out of the gel onto a membrane where it sticks and where its distribution according to size is retained.
5. Incubate the membrane in a solution containing a *labeled probe* so that it can hybridize to its complementary sequence on the membrane.
6. Wash the membrane under high stringency conditions where only the perfect hybrid duplex is stable. Detect the hybridized probe (fluorescence or radiography etc.) to determine the presence or absence of the DNA in the complex mixture of DNA on the membrane.



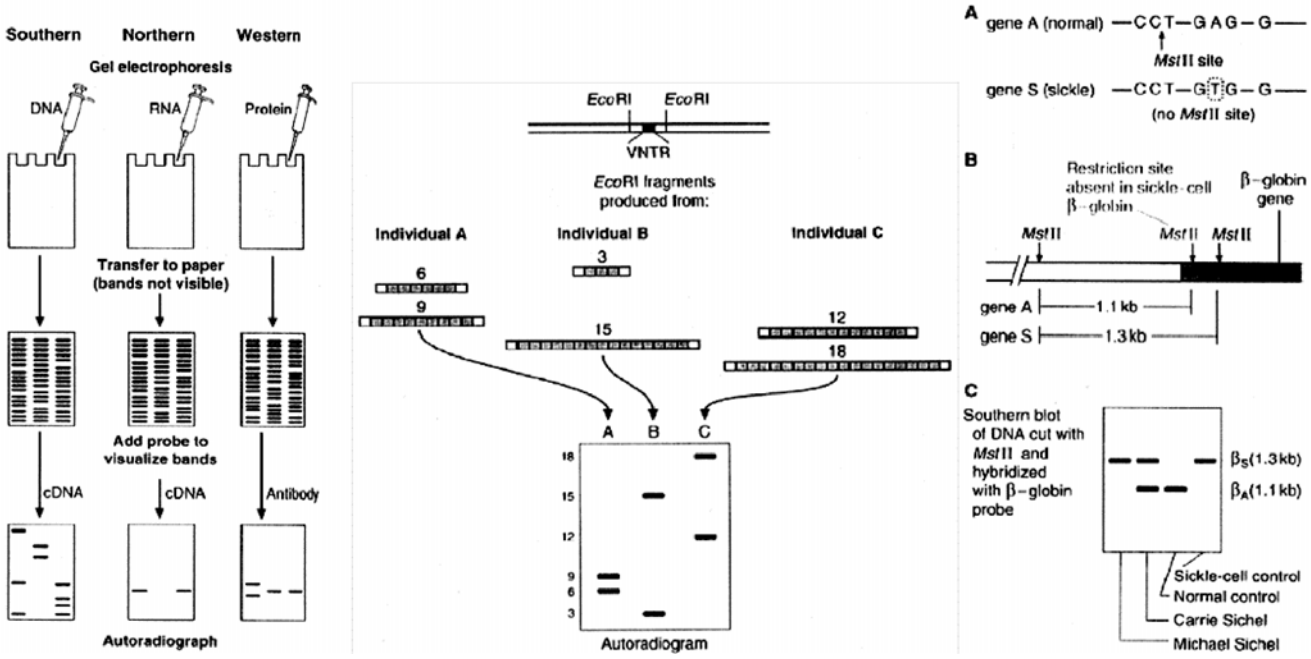
B. **Comparison of Southern, Northern and Western blotting.** See left figure on next page.

1. Southern – DNA is electrophoresed and transferred. The probe can be DNA or RNA.
2. Northern – RNA is electrophoresed and transferred. The probe can be DNA or RNA.
3. Western – Protein is electrophoresed and transferred. The probe is an antibody.

V. Polymorphism detection using Southern blotting.

A. **RFLPs** – Restriction Fragment Length Polymorphisms. A RFLP is a polymorphism that can be visualized because it changes the size of a restriction fragment. A RFLP can result from a deletion or insertion of DNA between two restriction sites. A deletion of the restriction site itself can also cause a RFLP. A RFLP can also be due to a single nucleotide change (i.e., a Single Nucleotide Polymorphism - SNP) that creates or destroys a restriction enzyme recognition site. See figure middle & right.

B. **VNTRs** – Variable Number of Tandem Repeats. VNTRs are a form of RFLP. There are many regions of human genomic DNA that contain nucleotide repeats. The number of repeats in each region varies greatly among individuals. A restriction enzyme that flanks a variable repeat region will produce a restriction fragment having a length that is proportional to the number of repeats. By assaying a battery of repeat regions, the pattern of the sizes of the repeats is used to uniquely identify a person. See figure on previous page. VNTRs are also assayed using PCR, which eliminates the need to transfer of the DNA to a membrane and probing (see below).



VI. Sanger dideoxy chain termination method for determining the exact sequence of nucleotides of a DNA region.

A. The *sequencing reaction* is similar to PCR described above. DNA polymerase catalyzes the addition of *deoxy*nucleotides to the 3' hydroxyl end of the oligonucleotide "primer" and to each successive nucleotide in the growing strand. In PCR, two primers are used, but in the sequencing reaction, only a single primer is used.

B. In contrast to PCR, *sequencing reactions* contain *dideoxy*nucleotides (ddATP, ddTTP, ddCTP, ddGTP) in addition to the four *deoxy*nucleotides (dATP, dTTP, dCTP, dGTP). The *dideoxy*nucleotides are *deoxy* at the 2' position (lack a hydroxyl group), as are the *deoxy*nucleotides; in addition, they are *deoxy* at the

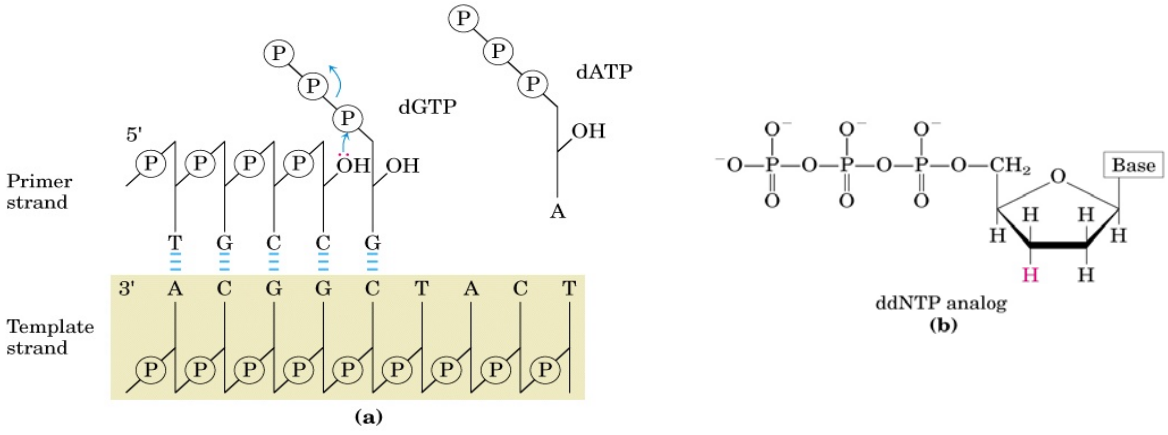
3' position. *Without the 3' hydroxyl, they cannot form a phosphodiester bond with the next incoming nucleotide*, and therefore the growing chain terminates at the point where a *dideoxy*nucleotide incorporates. See figure below.

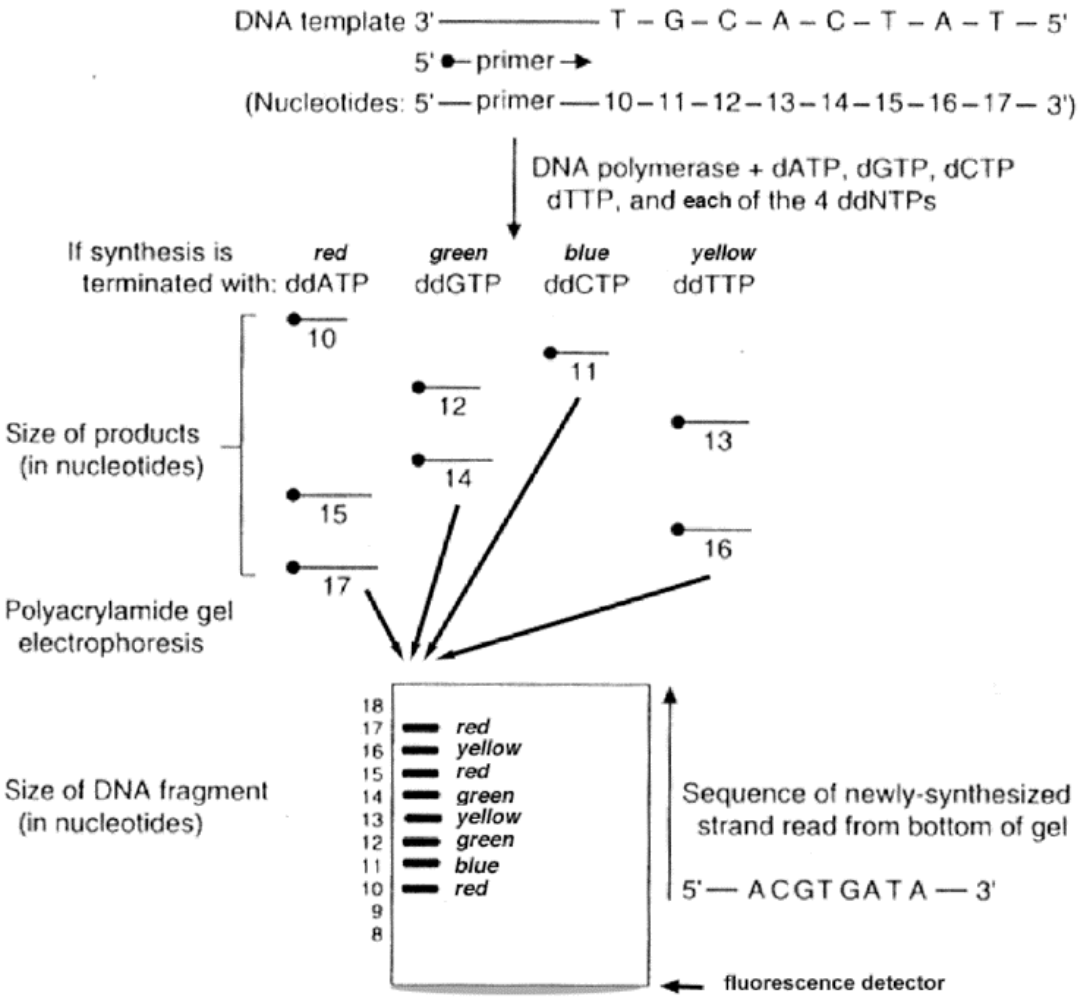
C. The *dideoxy*nucleotide triphosphates are each labeled with a different colored fluorescent dye (e.g., red-ddATP, green-ddGTP, blue-ddCTP, yellow-ddTTP). Thus, incorporation of the *dideoxy* into the growing chain terminates the chain and labels it with a color corresponding to the base at the site of incorporation. For example, if the fragment that is 50 nucleotides long is blue (cytosine), then the 50th nucleotide of the *template* is guanosine.

D. The concentration of *dideoxy*nucleotides is set low relative to deoxynucleotides so that a *dideoxy*nucleotide is incorporated randomly and only occasionally. The resulting infrequent and random termination permits synthesis of approximately the same number of each sized *dideoxy*nucleotide-terminated fragment (i.e., if the *dideoxy*nucleotide concentration were high, then the long fragments would be underrepresented due to early termination of the majority of the population).

E. The *sequencing reaction products* include *every possible* DNA fragment ranging in size from the "primer + 1" to the length of the complete template. The full size range of "colored" DNA fragments are electrophoresed and their color is read by a fluorescence detector as they electrophorese past it. Small fragments migrate fastest. For example, if the DNA fragments 10-17 nucleotides long are red-blue-green-yellow-green-red-yellow-red (in order from shortest to longest) then the sequence immediately downstream of the primer that is *complementary to the template* is ACGTGATA (see figure on next page, adapted from Marks Fig. 16.7) and the *template sequence* is 3'-TGCACTAT-5'.

F.




DNA template 3' — T – G – C – A – C – T – A – T — 5'
5' ● – primer →
(Nucleotides: 5' — primer — 10 – 11 – 12 – 13 – 14 – 15 – 16 – 17 — 3')
DNA polymerase + dATP, dGTP, dCTP
dTTP, and each of the 4 ddNTPs
If synthesis is terminated with:
red ddATP
green ddGTP
blue ddCTP
yellow ddTTP
Size of products (in nucleotides)
10
11
12
13
14
15
16
17
Polyacrylamide gel electrophoresis
Size of DNA fragment (in nucleotides)
18
17 red
16 yellow
15 red
14 green
13 yellow
12 green
11 blue
10 red
9
8
Sequence of newly-synthesized strand read from bottom of gel
5' — A C G T G A T A — 3'
fluorescence detector

G. **Clinical Correlation:** The antiviral drug AZT used for the treatment of AIDS is a thymidine analog that is a dideoxy chain terminator. It is not phosphorylated until it enters cells. Phosphorylation, which gives AZT negative charge, prevents it from diffusing across the cytoplasmic membrane and out of the cell. AZT is incorporated into the reverse transcript of the viral genome by *reverse transcriptase.* Chain termination of the viral DNA aborts the viral replication cycle. AZT is selective for reverse transcriptase, but some toxicity can result from incorporation of AZT into cellular nucleic acids by cellular polymerases.